WHAT IS CLAIMED IS:

1. A set of regulatory SNP polynucleotides or a set of polynucleotides complementary thereto, the set of regulatory SNP polynucleotides comprising a plurality of polynucleotides of at least 6 contiguous nucleotides, each polynucleotide of the set containing a regulatory SNP with 5', 3' or both 5' and 3' genomic flanking sequence, wherein the set of regulatory SNP polynucleotides contains a plurality of regulatory SNPs which collectively map to a plurality of TFC sequences such that each SNP lies within a TFC sequence and a genomic nucleic acid sequence from 30 nucleotides 5' to 30 nucleotides 3' to said each SNP is identical or complementary except for the SNP, to a portion of a genomic nucleic acid sequence from 30 nucleotides 5' to 30 nucleotides 3' to the TFC sequence.

2. A set according to claim 1 wherein the set of regulatory SNP polynucleotides comprises a plurality of polynucleotides of at least 10 contiguous nucleotides.

3. A set according to claim 1 wherein the TFC sequences are no longer than about 10 kb in length.

4. A set according to claim 1 wherein each of the TFC sequences comprise a sequence containing at least two transcription factor binding sites wherein likelihood of occurrence of the at least two transcription factor binding sites in the TFC sequence exceeds likelihood of occurrence in a random sequence.

5. A set according to claim 1 wherein each regulatory SNP when present in a genomic sequence, is associated with altered expression of a gene product as compared to

expression of the same gene product from a genomic sequence containing a corresponding normal allele instead of the SNP.

6. A set according to claim 1 wherein each regulatory SNP has a higher likelihood of occurrence in a population of individuals having a disease or predisposition to having a disease as compared to occurrence in a population of individuals not having the disease or predisposition to having the disease.

7. A set according to claim 5 wherein each regulatory SNP has a higher likelihood of occurrence in a population of individuals having a disease or predisposition to having a disease as compared to occurrence in a population of individuals not having the disease or predisposition to having the disease.

8. A set according to claim 1 wherein the regulatory SNPs collectively map to a plurality of transcription factor binding site sequences, wherein each regulatory SNP lies within a transcription factor binding site sequence and a genomic sequence from 30 nucleotides 5' to 30 nucleotides 3' of said each SNP is identical or complementary except for the SNP, to a portion of a genomic nucleic acid sequence from 30 nucleotides 5' to 30 nucleotides 3' to said transcription factor binding site sequence.

9. A set according to claim 8 wherein each regulatory SNP when present in a genomic sequence, is associated with altered expression of a gene product as compared to expression of the same gene product from a genomic sequence containing a corresponding normal allele instead of the SNP.

10. A set according to claim 8 wherein each regulatory SNP has a higher likelihood of occurrence in a population of individuals having a disease or predisposition

to having a disease as compared to occurrence in a population of individuals not having the disease or predisposition to having the disease.

11. A set according to claim 9 wherein each regulatory SNP has a higher likelihood of occurrence in a population of individuals having a disease or predisposition to having a disease as compared to occurrence in a population of individuals not having the disease or predisposition to having the disease.

12. A set according to claim 1 wherein each SNP lies within a genomic sequence from 100 nucleotides 5' to 100 nucleotides 3' to a TFC sequence and a genomic sequence from 30 nucleotides 5' to 30 nucleotides 3' of said each SNP is identical or complementary except for the SNP, to a portion of a genomic nucleic acid sequence from 130 nucleotides 5' to 130 nucleotides 3' to said at least one TFC sequence.

13. A set of probes or primers each of which comprises a polynucleotide of the set of regulatory SNP polynucleotides of claim 1 or the set of polynucleotides complementary thereto.

14. A set of probes or primers according to claim 13 wherein each probe consists of about 15 to about 30 contiguous nucleotides.

15. A biochip comprising the probes or primers of claim 13.

16. A set of probes or primers each of which comprises a polynucleotide of the set of regulatory SNP polynucleotides of claim 6 or the set of polynucleotides complementary thereto.

17. A set of probes or primers according to claim 16 wherein each probe consists of about 15 to about 30 contiguous nucleotides.

18. A biochip comprising the probes or primers of claim 16.

19. A set of probes or primers each of which comprises a polynucleotide of the set of regulatory SNP polynucleotides of claim 8 or the set of polynucleotides complementary thereto.

20. A set of probes according to claim 19 wherein each probe consists of about 15 to about 30 contiguous nucleotides.

21. A biochip comprising the probes of claim 19.

22. A set of probes or primers each of which comprises a polynucleotide of the set of regulatory SNP polynucleotides of claim 10 or the set of polynucleotides complementary thereto.

23. A set of probes according to claim 22 wherein each probe consists of about 15 to about 30 contiguous nucleotides.

24. A biochip comprising the probes of claim 22.

25. A method for diagnosing presence of a disease or predisposition for developing a disease in an individual, the method comprising determining whether a nucleic acid sample from the individual contains two or more regulatory SNP polynucleotides of a set of claim 5, wherein presence of the two or more regulatory SNP polynucleotides indicates presence of a disease or predisposition for developing a disease.

26. A method according to claim 25 wherein determining comprises determining whether the two or more regulatory SNP polynucleotides hybridize to the nucleic acid sample from the individual.

27. A method according to claim 26 wherein the two or more set of regulatory SNP polynucleotides are on a biochip.

28. A method for identifying a substance for treating a disease comprising testing a candidate compound for activity in modulating gene-product expression associated with a regulatory SNP polynucleotide of a set of regulatory SNP polynucleotides of claim 11.

29. A method of claim 28 wherein the gene product is overexpressed and the candidate compound is an antisense nucleic acid molecule comprising a complement of a genomic sequence which comprises the regulatory SNP polynucleotide.

30. A method of claim 28 wherein the gene product is underexpressed, wherein the regulatory SNP maps to a transcription factor binding site and wherein the candidate compound is a recombinant nucleic acid molecule which encodes a transcription factor which binds to the transcription factor binding site.

31. A method for treating or preventing a disease associated with a regulatory SNP in an individual, the method comprising administering to the individual a substance which modulates gene-product expression associated with a regulatory SNP polynucleotide of the set of regulatory SNP polynucleotide of claim 11.

32. A method of claim 31 wherein the gene product is overexpressed and the substance is an antisense nucleic acid molecule comprising a complement of the regulatory SNP polynucleotide.

33. A method of claim 31 wherein the gene product is underexpressed, wherein the regulatory SNP maps to a transcription factor binding site and wherein the substance is a recombinant nucleic acid molecule which encodes a transcription factor which binds to the transcription factor binding site.

34. A method of identifying regulatory SNPs, the method comprising determining that a SNP is a regulatory SNP if the SNP maps to a TFC wherein the SNP lies within a TFC sequence and a genomic nucleic acid sequence from at least 20 nucleotides 5' to at least 20 nucleotides 3' to said candidate SNP is identical or complementary except for the SNP, to a portion of a genomic nucleic acid sequence from at least 20 nucleotides 5' to at least 20 nucleotides 3' to the TFC sequence.

35. A method according to claim 34 wherein determining comprises determining that a SNP is a regulatory SNP if the SNP lies within a TFC sequence and a genomic nucleic acid sequence from at least 30 nucleotides 5' to at least 30 nucleotides 3' to said candidate SNP is identical or complementary except for the SNP, to a portion of a genomic nucleic acid sequence from at least 30 nucleotides 5' to at least 30 nucleotides 3' to the TFC sequence.

36. A method according to claim 34 wherein the TFC sequence is no longer than about 10 kb in length.

37. A method according to claim 34 wherein the TFC sequence contains at least two transcription factor binding sites wherein the likelihood of occurrence of the at

least two transcription factor binding sites in the TFC sequence exceeds likelihood of occurrence in a random sequence.

38. A method according to claim 34 wherein the regulatory SNP when present in a genomic sequence, is associated with an altered expression of a gene product as compared to expression of the same gene product from a genomic sequence containing a corresponding normal allele instead of the SNP.

39. A method according to claim 34 wherein the regulatory SNP has a higher likelihood of occurrence in a population of individuals having a disease or predisposition to having a disease as compared to occurrence in a population of individuals not having the disease or predisposition to having the disease.

40. A method according to claim 37 wherein the regulatory SNP has a higher likelihood of occurrence in a population of individuals having a disease or predisposition to having a disease as compared to occurrence a population of individuals not having the disease or predisposition to having the disease.

41. A method according to claim 34 wherein determining comprises determining that a SNP is a regulatory SNP if the SNP maps to a transcription factor binding site, wherein the SNP lies within a transcription factor binding site sequence and a genomic sequence from 20 nucleotides 5' to 20 nucleotides 3' of said each SNP is identical or complementary except for the SNP, to a portion of a genomic nucleic acid sequence from 20 nucleotides 5' to 20 nucleotides 3' to said transcription factor binding site sequence.

42. A method according to claim 34 wherein determining comprises determining that a SNP is a regulatory SNP if the SNP maps to a transcription factor

binding site, wherein the SNP lies within a transcription factor binding site sequence and a genomic sequence from 30 nucleotides 5' to 30 nucleotides 3' of said each SNP is identical or complementary except for the SNP, to a portion of a genomic nucleic acid sequence from 30 nucleotides 5' to 30 nucleotides 3' to said transcription factor binding site sequence.

43. A method according to claim 42 wherein the regulatory SNP is associated with an altered expression of a gene product as compared to expression of the same gene product from a genomic sequence containing a corresponding normal allele.

44. A method according to claim 41 wherein the regulatory SNP has a higher likelihood of occurrence in a population of individuals having a disease or predisposition to having a disease as compared to occurrence in a population of individuals not having the disease or predisposition to having the disease.

45. A method according to claim 43 wherein the regulatory SNP has a higher likelihood of occurrence in a population of individuals having a disease or predisposition to having a disease as compared to occurrence a population of individuals not having the disease or predisposition to having the disease.

46. A method according to claim 34 wherein determining comprises determining that a SNP is a regulatory SNP if the SNP lies within a genomic sequence from 100 nucleotides 5' to 100 nucleotides 3' to a TFC sequence and a genomic sequence from 30 nucleotides 5' to 30 nucleotides 3' of the SNP is identical or complementary except for the SNP, to a portion of a genomic nucleic acid sequence from 130 nucleotides 5' to 130 nucleotides 3' to the TFC sequence.

47. A method of generating a set of probes for use in detecting presence of regulatory SNPs, the method comprising identifying a set two or more regulatory SNPs which map to TFCs in accordance with a method of claim 34, identifying 5' genomic flanking sequence, 3' genomic flanking sequence or both 5' and 3' genomic flanking sequence, and preparing said probes.

48. A method according to claim 47 wherein each probe consists of about 15 to about 30 contiguous nucleotides.

49. A method of constructing a biochip for detecting presence of regulatory SNPs, the method comprising generating the probes of claim 47and placing the probes on the biochip.

50. A method of generating a set of probes for use in identifying regulatory SNPs, the method comprising identifying a set two or more regulatory SNPs which map to TFCs in accordance with claim 41, identifying 5' genomic flanking sequence, 3' genomic flanking sequence or both 5' and 3' genomic flanking sequence, and preparing said probes.

51. A method according to claim 50 wherein each probe consists of about 15 to about 30 contiguous nucleotides.

52. A method of constructing a biochip for detecting presence of regulatory SNPs, the method comprising generating the probes of claim 50 and placing the probes on the biochip.

53. A computer readable medium having a data structure for use in reporting regulatory SNPs, the data structure comprising a first field containing either or both of

sequence and genomic mapping location sequence information on SNPs; a second field containing either or both of sequence and genomic mapping location of TFCs; and a third field containing information on regulatory SNPs, the SNPs being identified by identities of either or both of sequences and genomic mapping locations of SNPs and TFCs.

54. A computer readable medium according to claim 53, wherein the first field containing nucleic acid sequence information on nucleic acid molecules which contain a SNP; a second field containing nucleic acid sequence information on TFC sequences; and a third field containing information on regulatory SNPs, the regulatory SNPs being identified by sequence identities of nucleic acid sequences containing SNPs and TFC sequences.

55. A computer readable medium of claim 53 wherein the nucleic acid sequence information on TFCs comprises nucleic acid sequence information on transcription factor binding sites and wherein the information on regulatory SNPs comprises identification of the regulatory SNPs by sequence identity of nucleic acid sequences containing SNPs and transcription factor binding site sequences.

56. A computer readable medium of claim 52 wherein the third field comprises information identifying the regulatory SNP and flanking sequence, the regulatory SNP being within a TFC sequence and a genomic nucleic acid sequence from 30 nucleotides 5' to 30 nucleotides 3' to said the SNP being identical except for the SNP, to a portion of a genomic nucleic acid sequence from 30 nucleotides 5' to 30 nucleotides 3' to the TFC sequence.